REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
09/08/2018	poster	08/09-12/2018

4. TITLE AND SUBTITLE
Non-Identifiable Coding for Longitudinal Research

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Jones, John R Capt

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
59th Clinical Research Division
1100 Willford Hall Loop, Bldg 4430
JBSA-Lackland, TX 78236-9908
210-292-7141

8. PERFORMING ORGANIZATION REPORT NUMBER
18020

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
59th Clinical Research Division
1100 Willford Hall Loop, Bldg 4430
JBSA-Lackland, TX 78236-9908
210-292-7141

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release. Distribution is unlimited.

13. SUPPLEMENTARY NOTES
Annual conference of the American Psychological Association, San Fransisco, AUG 9-12 2018

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Clarice Longoria
					19b. TELEPHONE NUMBER *(Include area code)* 210-292-7141

Non-Identifiable Coding for Longitudinal Research

Jones, J.R., M.A., Grabner, S.S., M.S., Gaschk, M., M.A., Kepple, J., M.A., Wolfe-Clark, A.L., M.S., Kalpinski, R., Ph.D., Bonilla, G., MPH, Talcott, W., Ph.D., ABPP, & Little, M., Ph.D., MPH
Wilford Hall Ambulatory Surgical Center

Background and Aims

Traditional methods of anonymous reporting garners vital data about otherwise sensitive information, such as risky drinking behavior. However, anonymous reporting does not allow longitudinal tracking. The use of self-generated identification codes (SGIC) may be useful in addressing this limitation. Researchers could use SGIC with military populations to infer the changes of individual alcohol use pre- and post-deployment without participants limiting their disclosures due to fear of self-incrimination.

However, many studies have found poor match rates between pre-test and posttest SGIC due to poor respondent reliability across time (for review see Schnell et al., 2010). The current study refined SGIC prompts and applied a statistically guided matching algorithm (Schnell et al., 2010; e.g., Levenshtein distance function method). The number of questions and SGIC content were selected based on stability over time and ability to discriminate between participants.

Methods

Active duty Airmen* (n=363) in technical school training were solicited to participate in the study during alcohol informational briefs. Airmen were informed they would be designing a SGIC to help gather sensitive healthcare information (e.g., drinking behaviors) anonymously over time. Upon consent, Airmen were given the non-identifiable coding questions and the AUDIT-C screener.

Following a 45-minute alcohol intervention, Airmen were re-administered the non-identifiable coding questions, prompting recreations of their original SGIC. Airmen were also given questions regarding their beliefs about the anonymity of responses. Pre- and post-intervention codes were then matched using the Levenshtein distance function method.

Only 309 participants completed pretests, therefore analyses were conducted in an effort to match these cases to post-tests.

Levenshtein Model

- Attempts to match pre- and post-intervention codes using the least number of operations changes between strings of data.
- Operation changes include: Insertion, Deletion, and Substitution.
- Example:

Original	Recreation 1	Recreation 2
B	B	B
6		
R	R	R
8	8	

- Recreation 1 would be matched with the original string due to a lower number of operation changes (1 change) compared to Recreation 2 (2 changes).

Questions used to generate code

1. First 2 letters of your mother's maiden name
2. The 2 letter abbreviation for the state in which you were born
3. First 2 letters of the name of the school where you began 9th grade
4. Day of the month were you born on
5. Number of **older** siblings do you have, alive and deceased
6. **Natural** hair color
7. First 2 letters of the mascot associated with your favorite professional sports team
8. First 2 letters of the make of your first car
9. Last number of your first home address
10. Number of siblings your mother has, alive and deceased

References

Schnell, R., Bachteler, T., & Reiher, J. (2010). Improving the use of self-generated identification codes. *Evaluation Review, 34*(5), 391-418.

*Airmen did not receive compensation for participation.

Approved for Public Release
Distribution is Unlimited

Results

- Using the Levenshtein method, 98.7% match rate was achieved (up to 4 changes).
- There was a 66% exact match rate (0 changes).
- Airmen were highly confident their responses were anonymous.
- Airman reported low worry about career impact when disclosing drinking rates during the study.

Discussion

- Matching algorithms like the Levenshtein distance function method are a viable method for short-term longitudinal pre-post test design research.
- Using matching algorithms with non-identifiable coding may serve to prompt more honesty from active duty service members when attempting to gather sensitive healthcare related data.
- Future study within the military context should attempt these methods in long-term longitudinal studies.

Disclaimer